

学校教育、劳动力质量和各国经济增长

埃里克·A·汉努舍克和丹尼斯·D·金科 著 陈大冰 译

原载:《美国经济评论》2000 年第 90 卷,第 5 期

经济增长率国际差异的近期理论分析将注意力集中在人力资本的作用。当前,大部份对长期经济增长的跨国间经验研究包括若干人力资本的代理变量;并且,这些变量无一不具统计显著性。然而,由于受到数据的限制,对于这些变量的规定,不得不做若干妥协。与工资决定的分析相平行,对人力资本的经验分析实际上通常使用若干手头可以检得的、关于正规教育的数量指标去反映人力资本。但这种做法看来并非妥当。本文所做的经济增长率的国际差异分析认为,劳动力适用的人力资本的一个基本组元是数学和科学技能(以下简称数科技能。——译者)。正规学校教育数量的指标或者投放在正规学校资源的指标并不能充当一类群体认知技能的适当的代理变量。不能忽视劳动力在质量上存在着显著差异。这一点一定会改进我们阐释国际经济增长率的能力。

在考察人力资本对经济增长的效应时出现两个问题:应该如何界定人力资本与经济增长的关系,应该如何计量人力资本。后一个问题是本文的重点。本文并不考虑另外一套带基础性的增长关系的建模方法。相反地,本文只是将好多人(譬如说 R·R Nelson, E·Phelos, P·Romer, S·Rebelo)从理论上开发出来的内生变量模型的直接应用。在这类最简单的建模方法中,观念和创造力影响经济增长率;而观念和创造力又通过研究开发活动或者通过采纳行为而与人力资本流量相关联。这类建模公式不仅表明为什么在某国的人力资本较多时该国产出水平较高;而且也表明人力资本较多时经济增长率也较高。

前期的增长研究集中在那些充当人力资本适用的代理变量的正规教育活动的指标。最常用的正规教育活动指标是小学在校学生率或者中学在校学生率。譬如说,在 P·Romer, R·J·Barro, N·G·Mankiw 等人的论文中就采用这两个指标;而在 R·Levine, D·Reneit 和 J·Zervos 的影响性敏感度研究中也突出这两个指标。但是,这类正规学校教育的流量变量要么不能精确地代表劳动力适用的人力资本存量;要么(甚至是)不能精确地代表教育变迁时期和人口迁移时期的人力资本存量变化。在处理这些问题时,Barro 和 J·Lee 率先采用个别国家的调整数据和人口普查数据开发出更优异的人力资本存量变量。但是,采用了这类变量后又冒出了一个富有挑战性问题。这个挑战性问题来自学校教育质量缺乏调整性。譬如说,难得有人相信,在美国中学修业一年等价于在埃及中学修业一年。实际上,Barro 曾经尝试将实际教育差异作为跨国间教育质量差异的粗放指标而列入他的经济增长回归模型。他发现,1960 年小学学生-教员比率和国家经济增长呈负相关关系;而中学学生-教员比率却和国家经济增长呈正相关关系,但其不相关程度却具非统计显著性。这一点使人难以相信,这两种相关关系的情况恰恰表明中学和小学具有教育质量差异。

在许多人力资本和经济增长模型的构建过程中也存在着概念性问题。这是因为源于人力资本增长的后续性经济增长常常要求人力资本的后续性增长。然而,由于某些简单的投资理由,人们并不期望学校教育受业年数以不受约束的形式扩展。不管如何,如果用认知技能和人力资本质量来表述,那么,持续性的人力资本质量增长就更具本质性,从而也让人能更顺当地去阐释人力资本支持经济增长模型。

本文直接地提出人力资本质量的测度问题。我们与其将本文重点放在学校教育投入的传统测度,倒不如去构造出那些基于学生认知成绩的新质量测度。这些认知成绩是从各种国际数科知识成就测验检得的。现在业经证明,按照这种方法测定的劳动力质量差异对经济增长率具有极强的效应。

我们搜集到 39 个国家(这些国家至少一次地参加过国际学生成就评鉴测验)的认知技能的直接观察值。虽然其中既有认知技能观察值,又有后续经济增长估计所需的经济绩效观测值的仅有 31 个国家,但我们却将认知技能和经济增长这两类质量测量数据增广到其他国家。增广的方法是从国际测验分数回归模型求得其他国家待补的遗缺数值。就直接测验数据子集和回归模型法增广的数据集合而论,关于劳动力质量和经济增长的相关性结论同样有效。即使学校教育几乎不呈统计显著性,但劳动力质量却呈统计显著性且正负符号无误。就 Levine 和 Renelt(《跨国间增长回归的敏感度分析》)所述的不受模型精确经验设定影响意义而言,劳动力质量观测值也是非常稳固的。此外,就解释国家的经济增长率高低位序排列而言,这些质量观测值也是非常重要的。

诸如劳动力质量和经济增长间相关的经验主义的观察免不了要听命于许多重要的、与不同经济体的时尚特征规范有关的问题和规定;也免不了要含糊其词地表述根本性的因果结构。既然处理这类因果关系问题的工具变量策略不切实际,我们只好追随三种不同的工具变量策略。这三种策略的前提都是:我们已经对经济增长因果结构做了重要因素的识别工作。第一种策略是,如果较强劲的经济增长导致国家对学校的投资,经济增长就会引致学生成就增大。但是,国际生产函数的直接观测并不支持这种策略;认为它并非因果效应途径。第二种策略是,如果国家的若干个未经观测的特征因素既影响学校教育绩效又影响国家经济其它部门绩效,那么,我们所观察的学校教育与国家经济增长的关系就是属于伪相关的。但是,美国按上学区划分的移民收入模型的数值估计和我们所做的劳动力质量估计值却表明:我们的劳动力质量估计值直接关联个人的劳动力技能和生产率。第三种策略是,这类回归检验可能恰恰做出这样确认:东亚国家的经济增长高;东亚国家的国际学生成就测验得分也高。然而,对那些不包括各种东亚国家子集的国家样本来说,劳动力质量对经济增长有影响效果还是成立的;尽管影响效力稍微差些。

在劳动力质量对经济增长的回归检验中,要留心受估计的劳动力质量对经济增长的影响程度。对移民微观生产率的估计值(劳动力质量差异对收入的影响估计值比按增长方程式求得的计算值会稍微高些)表明:潜在的劳动力质量对经济增长的因果效应程度具有不确定性。随着人们将生产率差异等级效应翻译为经济增长效率的不同,劳动力质量对收入效应的估计值会比按增长方程式计算出来的估计值明显小些(这表明更有可能存在其他省略因素)。不管如何,即使有数量级的不确定性,我们还是得出这样的结论:劳动力质量是直接与生产率和经济增长相关的。

1. 劳动力质量的测定 在考察人力资本时,一般说来,人力资本的数量描述不外取自两个来源:学校教育投入物(诸如教学费用或者教育薪金)或者个人认知技能的直接测度。采用认知技能直接测度的方法具有这样的优点:允许正规学校教育以外的因素所生成的劳动力质量差异。然而,只有在与人力资本有关的重要侧面受到认知测验的局部度量时,采用正规学校教育投入物测度才具有潜在性优点。我们既采用认知技能测度也采用正规学校教育投入物测度。因此,本文的分析核心不外是:根据一个与人力资本数集相一致的认知技能测验数值集合建立回归模型,然后使用该数集。

在比较跨国间认知成就时,我们利用在过去 30 年所做的六次自愿性国际数科成就测验的数据。在这六次测验中,国际教育成就计量学会(IEA)做了四次;国际教育进步评鉴组织

(IAEP)做了两次。由于成立于 1959 年,IEA 在几乎所有初等教育和中学教育方面的教育比较研究开发上长期扮演独特的角色。另一方面,自 1988 年以来,国际教育进步评鉴组织已着手逐步地将美国全国教育进步评鉴组织开发的统计技术和统计程序完备起来。美国全国教育进步评鉴组织是 1969 年以来美国全国主要教育进步测验机构。国际教育进步评鉴组织的测验已与美国教育测验接轨,但国际教育成就计量学会仍然将工作重点放在国际方面。这个学会的教育测验并不与任何特定国的教育测验挂钩。

将重点放在数科技能是与理论上强调研发活动作为增长资源的重要性相一致的。深刻领会到数学和科学方面的有才华学生是未来工程师和科学家的人才库。至少,就美国情况而言,J·Bishop 在《智能对工资、失业和工作成绩的影响》(1992 年)也单独地论证了数学和科学在确定个人的生产率和收入方面的重要性。另外,目前虽然还有其它学科的测验资料,但它们不能直接了当地用来同数学和科学分数相比较。因此,在本研究中,我们就不予采用了。

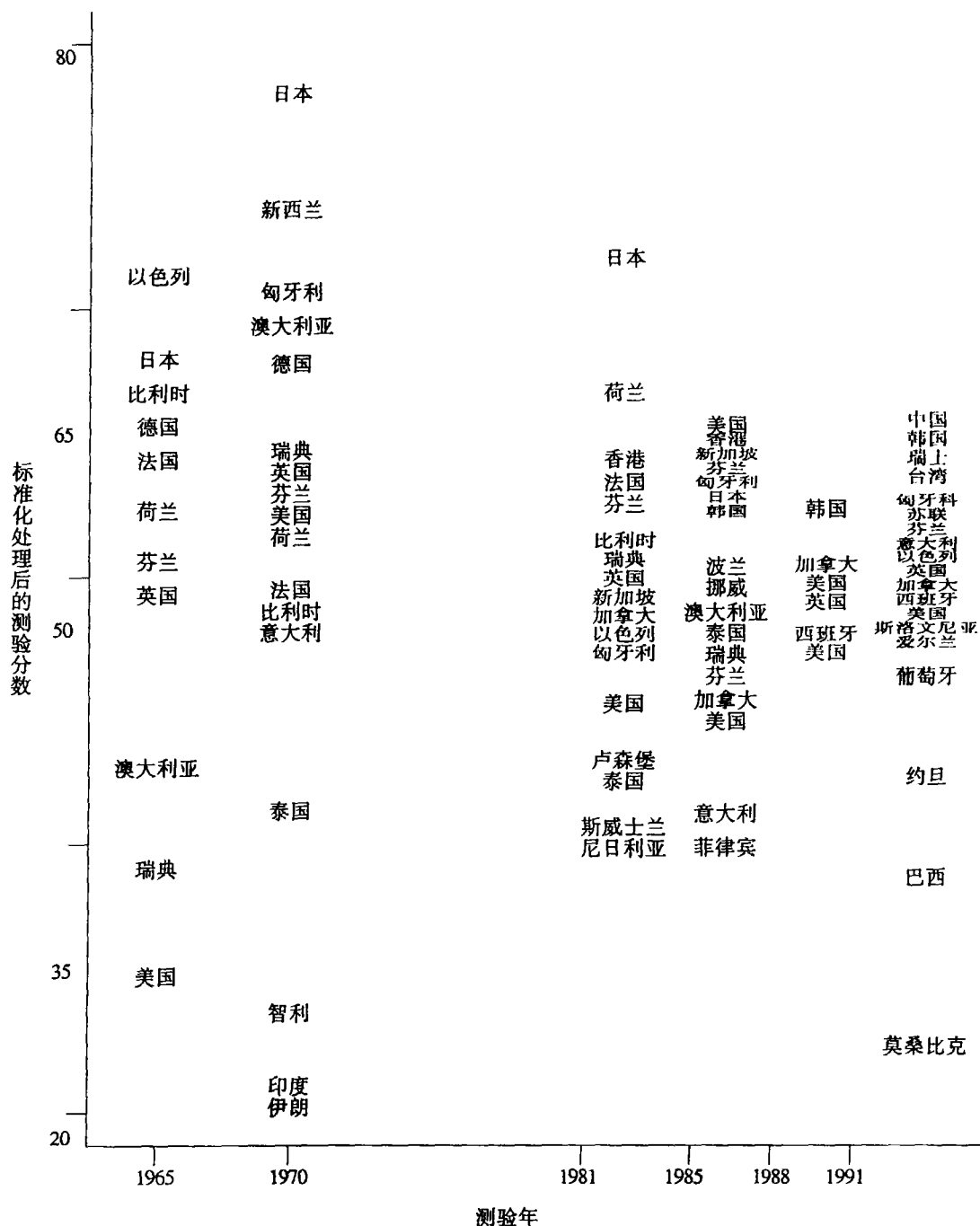
为了开发一个单一的劳动力质量测度,我们将所收集的 1999 年各国国际数科测验信息汇总起来。此国际数科测验并非直接测量特定存量的工人(他们在其一生的不同时间接受学校教育)的数科技能,虽然不同的国际测验的混合搭配与相应的时间阶段相差不远。尽管难以收集数据,但常规的假设却是,学校体制的质量随时间推移而逐渐演进。其所以如此,部份的原因是既与教育方法的稳定性有关,也同教员和其他教育人员的逐渐转换有关。尽管如此,这种汇总各技能测分以估计劳动力技能总体测分的方法并不再考虑在各个次阶段如何测定劳动力技能和经验增长的进展了。

我们用两种方法将可收集的各国逐一测定的分数汇总起来。26 个分数系列(这些系列反映不同阶段、子集分数和年份)是可用的(就各国的不同子集而言)。第一种汇总方法使用一种乘法转换运算。这种运算将每一个测得分数系列转换为某种平均值(此平均值 = 50)。这种乘法转换运算根据这样一种强假设:世界数科测分的各期平均值呈常数;受数科技能测验的国家依随机法从全球各国检得。第二种汇总方法是将来自全国教育进步评鉴组织的时间系列信息所提供的附加信息并合。在 1969 年至今的各个不同时间里,美国的 9 周岁学生、13 周岁学生和 17 周岁学生都受数科技能测验。这些测验是在与时序比较相一致的基础上进行的。因此,测验本身也提供一个国际测验的美国测分得以对准的测验基准点。这样,就允许每一个国际测验系列的测分平均值作出适当的偏移。偏移程度应该与美国全国教育进步评鉴组织的测分偏移和美国在每一次国际测验的中所得分数的平均值相应。这样构造出来的每一个国家学校教育质量测值是各个可收集到的、业经转换运算处理过的测验分数的加权平均值(其权数是国别标准误(σ))的(标准化的)倒数。

图 1 按时序将所收集到的测验分数标示出来。在图 1 中,对 6 个基础性国际测验评鉴中的每一个评鉴而言,我们将一国的每一年龄组和每一次测验的分数综合积算成一国的测验分数。(在评鉴综合积算时,我们将每一个国际评鉴的全世界各国分数平均值定为 50)然后,依据每个国家,我们将此质量测定分数逐年组合,从而得到单一的质量总值。

全国教育进步评鉴组织所计算的美国学生总分在 70 年代呈下跌趋势,在 80 年代呈上升趋势。此趋势模式如图 1 的美国各个国际测验得分分布。这样,我们的 QL1 和 QL2 的测定值呈高相关($r=0.92$)。图 1 示出 31 国用的 2 个构造出来的质量指标的总结性统计值。这 31 国也都有完整的下面增长分析所需用的经济数据。第 1 个系列(QL1)是基于 6 个测验用的世界平均分数(50)的;第 2 个系列是以全国教育计划评鉴组织的美国测验分数为基础的。

图1 1965年、1970年、1981年、1985年、1988年和1991年国际数学测验和科学测验的各国得分图



2. 质量对增长的效应 指导本研究工作的基础性建模思路是仿效内生性增长模型。按照内生性增长模型的思路,某国的增长率是直接与人资本存量相关的。种种内生性增长模型的建模方法都说明内生模型公式的经验变量的界定方法。P·Romer 的《内生性技术变化》(1990年)表明,人力资本影响观念和新技术的供给。在 Rebelo 的《长期政策分析和长期增长率》所述的 AK 模型中,增长直接与那种包含人力资本的总资本存量相关联;不过,这种人力资本具有某种不具备人力资本报酬递减性质的关键因素。这类模型的前身也正是 Nelson 和

Phelps 以及 F·Welch 所采用的技术模型。Nelson 和 Phelps 模型以及 Welch 模型表明,一国的人力资本影响某个该国能使其新技术的投放得当的新旧技术比率。从而,也影响该国的经济增长率。尽管 Nelson 和 Phelps 模型和 Welch 模型的政策含义不同,但两者均可引致本文所用的基本模型的建立。

当然,内生性增长模型的界定公式并不是用以界定增长差异方程的唯一模式。因此,围绕着建立最佳模型问题免不了有许多争论(关于选择增长模型的讨论,请参阅 Barro 和 Xavier - Sala - i - Martin 合著《经济增长》一书)。在这种讨论中,有一部份是以理论为基础的。一般说来,它们和这类增长模型的长期属性有关。这种讨论的另一部份是以经验为基础的。模型变量的经验规定的基本问题是:人力资本存量或者人力资本存量的变化应不应该参与增长率的决定机制。如果人们将教育视为一种直接性生产投入物,那么,增长率应该与不同的投入物增长有关;并且,人力资本存量变化应该是增长的适当的解释因子。我们对这类备择增长模型的各个方面做了检验。但是,由于检验结果随模型的变量规定和变量含义而异,这些检验并不具有结论性价值。过去对此问题的讨论大多集中在人力资本数量指标的规定和含义。在此问题上,中心命题是:要是诸如本研究所引发的质量关注是经济增长差异的关键因子的话,现存检验似乎就是一种误导性的检验。

表 1 构造劳动力质量测度用的总结性统计量(31 国)

劳动力质量测度	中位数	平均数	标准差	最小值	最大值
QL1	48.76	46.61	10.86	20.79	60.65
QL2	54.52	51.28	13.48	18.26	72.13

注:参阅 Hanushek 和 Kim 的《学校教育,劳动力质量和经济增长》(1995 年)。该文详述该文所收测验和劳动力质量测度的构造情况。QL1 设定六次(基础性)测验每一次测验的世界平均值等于 50;而以美国国际学生成绩的测验分数(此分数曾参照美国全国教育进步评鉴组织的全国时序分数模式加以修正)为基础的所有分数,则由 QL2 加以调整。参与 1 次或不止 1 次测验的国家有 39 个。不过,本表的数据国家仅限于那些除学生成绩数据外,还有用于后续增长分析所需的经济数据的国家。

表 2 1960~1990 年跨国间由劳动力质量
导致的增长模型基线估计量(因变量:按人均 GDP 计算的平均年增长率($\times 100$))[31]

	(1)	(2)	(3)	(4)	(5)	(6)
初始人均收入(1960 年)(单位:1000 美元)	-0.609 (0.186)	-0.472 (0.096)	-0.460 (0.103)	-0.745 (0.181)	-0.481 (0.093)	-0.517 (0.112)
学校教育数量(S)	0.548 (0.209)	0.103 (0.126)	0.100 (0.146)	0.519 (0.195)	0.106 (0.119)	0.116 (0.139)
年人口增长率(GPOP)				-0.713 (0.224)	-0.038 (0.215)	-0.250 (0.211)
劳动力质量(QL1)		0.134 (0.023)			0.133 (0.024)	
劳动力质量(QL2)			0.104 (0.015)			0.098 (0.015)
常数	2.265 (0.863)	-1.900 (1.004)	-0.989 (0.910)	4.092 (0.974)	-1.756 (1.346)	-0.151 (1.142)
R ²	0.33	0.73	0.68	0.41	0.73	0.69

注:系数下面括号内的数字是 Huber - White 估计标准误差。

本文所强调的是,内生性增长模型的不同经验版本的比较。在比较时,我们特别将注意力放在如何解释那些由于包括质量指标而受到影响的跨国间增长差异。因为只观察人力资本的截面差异,所以这类分析就无法提供其它各种内生性增长模型检验。

表2报告我们基线的跨国间回归(它描述了1960年至1990年间的平均实际人均GDP增长)的计算结果。在求此计算值时,我们先设定一小套决定增长率的因子;然后,研究劳动力质量影响的数量级和稳定性。接下来的估计步骤是:将考虑的范围扩展到其他国家以及其它受到人们经常估值的因子的潜在影响和令人困惑不解的影响。

用于数据完整的31个国家的最简单模型将经济增长和初始收入(Y60)以及增长和Barro-Lee学校数量(S)[纵栏(1)]关联起来。一条可备选用的基线使用初始收入(Y60)加Barro-Lee学校数量加人口年增长率($\times 100$)[GPOP][纵栏(4)](变量定义和数量来源可于附录A检得)。这些基线估计值给出相当常用的结果;并且,对31个数据完备的受观察国家组的经济绩效变异(33~41%)作出适当的解释。我们的基本结果与前人已往的估计值相一致。初始收入对增长有负面影响。这一点支持原先关于增长率条件收敛的设想。学校教育数量(S)对增长有正面的强影响。

表2纵栏(2)~(4)所载的我们添加的备择劳动力质量指标的相应估计值表明:劳动力质量和人均增长率有很强的相关性。增多劳动力质量的测定值的个数(不管是QL1或者QL2的个数)会使经调整的 R^2 值膨胀到0.7上下。此值比用更简洁的回归模型求得的大得多。劳动力质量(不管用QL1或者QL2测定)增加一个标准差,人均实际经济增长率大约增长1.4百分比点/年(譬如说, $0.134 \times 10.86 = 1.46$)。与此对比,学校教育数量增长一个标准差,人均实际经济增长率只增长1/4百分比点($0.10 \times 2.63 = 0.26$)。(注意,在包括直接学业成绩测度后,学校教育数量大幅度下降)。由于这些国家的无附加条件的增长率标准差为1.75%,所以,估计所得的劳动力质量效应的等级是很高的(关于劳动力质量效应的等级,我们下面还要回头来讨论)。

用加进人口增长这个因子来对基本模型作适当复制[栏(4)至栏(6)]。复制出的模型几乎没使学校教育质量的等级和相关显著性发生变化。人口增长对经济增长的影响估计值通常呈负值,但它们对模型变量的规定却甚敏感,并且,当考虑到劳动力质量时,这些估计值均异于零。再者,正如其他先前人们所估计的影响那样,人口增长和经济增长之间的因果性关系也是应予考虑的事项。这是因为较高的收入可能使某个国家的出生率下降。

基本模型只将注意力集中在学校数量和学校质量的效应上。然而,这类效应可能是补充性的。若干对线性互动项的实验使人难于相信那些对最高或最低学校数量样本和最高或最低学校质量(此质量业已量化)样本的计算结果。另一方面,如果不用基数模型、而改用对数模型,我们得到的是:和不带互动因子的基本加法模型一样的实际上能加以识别的质量结果。由于资料阙如,我们无法令人满意地对备择的泛函型增长模型进行研究。

基本结果给那种用数科认知成就度量的劳动力质量差异以强有力的支持。这些结果也提供研究基线。借此基线,我们既可以对劳动力质量的效应渠道作一分析;也可以对将该方法推广到其他国家的成果作一分析。

3. 因果关系 A篇:学校教育质量的决定因素 增长能使某个国家得到增量性资源。这些增量性资源的一部份又可倒过来溶入人力资本投资。因此,可以设想:先前估计的相关关系可能将较高劳动力质量的因果效应加以夸大。现在我们来考察下面这样一个模型描述:

$$(1) g_i = \mathbf{X}_i\beta + \gamma QL_i + \epsilon_i$$

$$(2) R_i = \mathbf{W}_i\delta + \eta g_i + u_i$$

$$(3) QL_i = \mathbf{Z}_i\alpha + \pi R_i + v_i$$

国家(i)的增长(g)由劳动力质量(QL)加由其他因素构成的向量(X)决定(方程式 1)。然而,和其他因素(W)一道,增长(g)也对决定那些用于学校和人力资源生产的资源总量(R_i)(方程式 2)作出贡献。这种建模彰显出这样一种事实:政府不能直接影响增长结果;相反地,政府必需采取不同的、基于学校组织和基础性生产函数的间接性政策。不管怎样,资源和其他投入物(Z)共同决定劳动力质量[方程式(3)]。对方程式(1)的简单估计并不能提供质量对增长的因果性影响的估计值。相反地,这种估计却反映增长对质量的影响。这种影响也体现在结构性参数 η 和参数 π 上。

诚然,要用手头的资料去估计上面 3 个方程式的参数是办不到的。我们不采用这种 3 个方程式联合求解方法。在此处,我们所采用的求解方法是直接对方程式(3)进行估计;具体说,就是直接对人力资本生产函数进行估计。此人力资本生产函数的功能是将待测定的劳动力质量联系跨国间各种学校和各种劳动素质的人口群体的总体资源(Z)。参数 ν 和参数 ν 不具相关关系。因此,估计方程式(3)就提供生产函数参数的一致估计值。这种估计是按照学校教育成绩的微观水平进行研究的。过去对这一类模型的估计(对美国和发展中国家的估计)无法找到学校资源和学生成绩间的一致关系。因此,没有必要将这种估计工作推广到国际比较。如果和国家与国家间的资源差异相比较,一国内部的资源差异就微不足道了。这一点表明:即使资源的边际差异不大,在跨国间发现的大等级差异也发生重要的后果效应。

我们假定,学生平均能力的国际水平在跨国间并无差异(至少,对我们此处所考虑的其它决定因子而言,此变量是外生变量)。这一点促使我们将注意力集中于标准的、手头收集到的资源测度和总群体的素质特征测度。就先前的增长模型而言,我们感兴趣的是:测定劳动力总体质量;从而,将各个不同的测验分数归结为用于每一个国家的单个(Single)的国家测验分数。然而,在本研究的模型中,我们感兴趣的却是具有基础性的生产关联性。根据这种关联性,我们就能够利用分类测验数据去探究学校资源和家庭资源是如何同特定观察人群组的学校成绩联系起来的。在估计学校生产关联性方面,我们在六种彼此分开的测验中,考察那些业经标准化的个别国家测验分数(即,在图 1 中标出的个别国家数据点)。然后,将这些分数和相应的观察人群组专用的家庭特征值和学校资源关联起来(参阅附录 B)。这种方法给出高达 70 国的测验组观察值和观察必备的学校投入物数据。这种方法还保证:对学校教育成就说来,学校投入物变量必定是外生变量。

表 3 列出生产模型的若干变异模型。这些模型的总细节是,学校资源的差异对学生测验分数并没有强影响。不同资源测度对学生测验分数的效应估计值要么不具统计显著性,要么(在更常见的情况下)虽具统计显著性,但其正负符号却与期望相违背。不管是哪种学校资源测度(教员-学生比率、每个学生的经常性开支、每个学生的总开支或者其它许多测度),上面这种发现尚不失效。用成年人群体学校教育数量作为其代理变量,父母教育变量倾向于呈正显著性(就传统的统计检验水平而论)。再者,较高人口增长率也倾向于具备较低的学校教育成就。这一点同关于儿童素质和儿童人数与大家庭规模影响间的衡平的标准命题并不相悖。本研究的表 3 未列出的其它备择模型包括世界地区虚拟变量。但是,在这些备择模型中,上述受测学校资源对学生成就的效应没有什么变化。极其凸显的是,学生-教员比率的反常效应并不是那种我们在许多东亚国家所发现的上大班的授课效应的简单因果推理。

最后,测验结构本身可能影响受观察学生成绩模式。入学人数的变异可能导致那种因纯粹选择效应而生的分数变异。这是因为就其典型性而言,较低的学校完备率的国家可能是对总年龄组群体中具有较大选择性的人口份额进行教育成就测验的。表 3 所列的模型的变量业经再度估值。在估值前,人们先加入 5 年期(包括特定测验)的中等学校入学率(表中未列出)。中等学校入学率一致地虽呈正符号但不具统计显著性。(这一点恰恰与人们预期的重要选择

效应的相关性相反)。不过,学校资源效应模式却受不到什么影响。

表3 数科成就的测验成绩:国别分数与组别分数(因变量:测验年 t 的标准化成绩)

	(1)	(2)	(3)	(4)	(5)	(6)
IEA ^① 数学1				44.09 (5.16)	50.15 (5.61)	40.78 (7.03)
IEA ^① 科学1				46.44 (5.87)	54.30 (5.48)	41.98 (7.01)
IEA ^① 数学2				48.49 (6.04)	55.49 (5.65)	41.30 (7.11)
IEA ^① 科学2				45.94 (6.35)	51.66 (6.16)	37.97 (7.62)
IAEPA ^② 数学科学				47.14 (5.47)	52.45 (4.91)	38.24 (6.57)
成年学校教育(S_{t-1})	2.04 (0.82)	1.62 (0.76)	1.54 (0.64)	2.70 (0.70)	1.75 (0.73)	1.59 (0.64)
小学学生-教员比率($PT - pri_{t-1}$)			0.066 (0.16)			0.09 (0.15)
当年每名学生的公共支出(PPE_{t-1})	-0.69 (0.19)			-0.766 (0.21)		
教育总支出/GDP($EXPEND_{t-1}$)		-165.90 (90.66)			-189.78 (88.69)	
年人口增长($GPOP_{t-1}$)	-4.65 (1.68)	-4.60 (1.36)	-2.64 (1.96)	-4.86 (1.94)	-4.98 (1.42)	-2.81 (1.91)
常数	46.46 (5.17)	52.27 (4.94)	40.80 (6.55)			
国家数	69	67	70	69	67	70
R^2 (调整后的 R^2)	0.25	0.19	0.25	0.22	0.26	0.25

注:① 国际教育成就计量学会(IEA)。

② 国际教育进步评鉴组织。

本样本由各参与国依一国一观察值组成。就六个测验年的每一个测验年而论,不同测验年的参与国国家数虽不同,但每一参与国均参与国际测验。每个测验年 t 的测验分数均经标准化处理。标准化的分数平均值等于50。系数下面括号内的数字是 Hober-White 标准误差。

与先前跨国间增长相关性分析有明显关联的是多头并举性(Simultaneity)的似然度:较快增长的国家似乎将更多资金投在扩充学校教育、厂房设备,等等。有人(譬如说,J·Mincer、Bils和Klenow)特别强调经济增长对人力资本开发的作用;他们并不强调人力资本对经济增长的影响。因为收入和支出并不系统地影响劳动力质量,所以在我们的增长模型中,我们对劳动力质量的因果说明就更有说服力了。之所以如此,其原因是: $\pi = 0$,从方程式(1)到方程式(3)的反馈回路也就消失了。

4. 国家扩充样本 为了进一步探索国家间的增长差异,我们将受分析国家组大大扩充。在分析这些国家时,我们借助一个基于受观察特征的劳动力质量因变量的模型方程式设计。我们的估计策略的焦点是上面所列的方程式(1)至方程式(3)。我们观察一组 n_1 个国家的 g , X , QL 和 Z 。与此同时,我们也观察另一组国家($n_1 + 1, \dots, n_2$)的 g , X 和 Z 。在开始时,我们从方程式(3)、对第1个 n_1 个观察样本国家组求取 α (α 是国家因子对受测劳动力质量的效应)的一致估计值。然后,我们使用此估计值($\hat{\alpha}$),采用联立求解法对第1个具有 n_1 个观察国家样本组求方程式的估计值;此后,对余留($n_2 - n_1$)个国家观察样本组求方程式(4)

(4) $g_i = X_i\beta + \gamma Z_i\hat{\alpha} + [\gamma u_i + \varepsilon_i + \gamma Z_i(\alpha - \hat{\alpha})]$ 的估计值。此处,括号内的项表明方程式(4)的复合误差项。

如果就具有完整数据的国家而言, γ 和 u 并非相关;那么,在很一般的条件下,上述方程式组可求得 γ 的一致估计值。但是,显然方程式(4)的完全误差项将有一个比方程式(1)误差项的方差大的方差。为了得到标准误的估计值,对异方差性的校正就显得重要了。如果异方差性的唯一根源是来自方程式(3)国家扩充组的 QL 的估计程序,那么,一个简单的、用于第 1 组 (n_1 个)观察样本的和用于第 2 组 ($n_2 - n_1$ 个)观察样本(第 $n+1$ 样本, ..., 第 n_2 个样本)的采用不同方差估计值的二阶段估计量就不能说是不恰当的了。然而,异方差性还有其它根源。因此,我们采用更一般、更稳固的估计技术。

在设计方程式(3)的因变量(劳动力质量)时,我们回到了简单的劳动力质量(QL₁ 和 QL₂)总量国家数值的截面样本。我们将这些总量值和在整个 1960 年~1990 年时期的平均人口和学校教育特征值关联起来。我们的第一个目的是,开发一个用于设计因变量的方程式,我们使用扩大模型(这些模型的扩大程度超过先前生产函数的估计值)。

表 4 用于劳动力质量的预测模型

	因变量 = QL1			因变量 = QL2		
	(1)	(2)	(3)	(4)	(5)	(6)
小学入学生率(ENROLL - pri)	81.50 (16.54)	85.78 (23.32)	86.23 (24.81)	62.36 (24.44)	73.28 (29.54)	75.32 (31.30)
学校教育数(S)	-0.714 (0.53)	0.038 (0.65)	0.320 (0.64)	1.651 (0.77)	0.97 (1.18)	1.181 (1.14)
小学学生 - 教员比率(PT - pri)	0.142 (0.06)			-0.043 (0.13)		
当年教育支出/GDP(RECUR)			55.49 (119.3)			133.23 (176.1)
教育总支出/GDP(EXPEND)		121.07 (109.9)			170.37 (169.1)	
年人口增长(GPOP)	-3.11 (1.76)	-3.08 (1.98)	-2.80 (2.06)	-4.105 (2.36)	-4.17 (2.65)	-3.85 (2.62)
虚拟变量:亚洲国家(=1)	5.123 (3.56)	7.52 (2.76)	6.33 (3.07)	12.25 (6.54)	13.77 (4.76)	12.76 (5.02)
虚拟变量:拉美国家(=1)	-4.004 (2.31)	-3.87 (2.93)	-4.31 (3.39)	-1.24 (3.689)	0.203 (4.07)	0.432 (4.53)
虚拟变量:非洲国家(=1)	3.170 (2.38)	2.94 (2.32)	3.30 (2.68)	11.97 (5.73)	8.71 (3.44)	9.16 (3.72)
常数	-35.59 (17.20)	-37.19 (22.41)	-36.19 (24.55)	-13.10 (25.76)	-28.40 (27.65)	-29.27 (30.17)
国家数	31	30	30	31	30	30
R ²	0.73	0.72	0.71	0.68	0.66	0.65

注:系数下面括号内的数字是 Huber - White 标准误差。

正如表 4 所示,一般说来,在确定劳动力质量差异方面有三个重要的基本因素。初级学校入学率强烈地影响学生成绩。这可能是因为每一个国家都很重视教育。人口增长率越高,劳动力质量就越低。撇开其他特征变化不谈,初级学校入学率的地区差异很显眼,这一特点,亚洲国家的表现最为突出。一般说来,平均学校教育数量和学生成绩之间呈正相关;虽然就典型

性而言,这种正相关并不具显著性(按传统统计检验水平而言);再者,学校资源和劳动力质量之间也不呈强相关。不管亚洲地区(就传统而言,该地区的学生-教员比率高,学生成绩也高)是否存在一个虚拟变量,但学生-教员比率变量都出现不正确的相关正符号。支出变量的估计值虽说是正符号,但并不呈统计显著性。

我们利用(表3)纵栏(2)和纵栏(5)去构造一个劳动力质量因变量($QL1^*$ 和 $QL2^*$)的扩大组。 $QL1^*$ 和 $QL2^*$ 将那些劳动力质量的测验值同所有虽不具有受测验的实际数据、但却具有方程式(3)右侧各自变量数据的劳动力质量的预期值结合起来。但是,特别令人注意的是,相对说来,参与测验的发展中国家寥寥无几。这些发展中国家就是位于样本分布中学生成绩最低的若干个观察样本。因为排在学生成绩最低的观察样本稀稀拉拉(参阅图1),我们模型的主要增长估计值迫使我们舍弃所有预期学生成绩分数低于20(20是实际观察分数的下限值)的国家。但是,对抽样准则说来,后续增长估计并不是很敏感的。再者,最终证明,在两个劳动力质量因变量之间也并没有什么差异。不过,在增广国家样本中,这两个劳动力质量因变量之间却有单相关(相关系数为0.95)。

作为这样模型方程式构造的一种独立校正,人们可以将劳动力质量的估计值关联到第3次国际数科学学习测验1995年8级国家分数。联合式第3次国际数科学学习测验分数与因变量 $QL1^*$ 和因变量 $QL2^*$ 的相关系数分别为0.67和0.66。如果将南非(该国的分数比32个业经匹配的国家样本中的次低分数低1个完全标准差)舍弃不计,那么这两个相关系数就分别升到0.76和0.78。重要的是,在第3次国际数科学学习测验样本中,与先前24个国家样本相反,新受测验的8个国家样本的此两相关叙列的相关性并没有显著差异。这种高相关系数也加强先前认为学校教育系统具有相对稳定性的假设。

表5 1960~1990年增广劳动力质量样本用的跨国间劳动力质量所影响的增长模型

	(1)	(2)	(3)	(4)	(5)	(6)
初始人均收入(Y60)(单位:1000美元)	-0.382 (0.081)	-0.390 (0.079)	-0.453 (0.078)	-0.370 (0.084)	-0.384 (0.082)	-0.442 (0.081)
学校教育数量(S)	0.127 (0.089)	0.117 (0.093)	0.112 (0.093)	0.120 (0.096)	0.103 (0.100)	0.112 (0.104)
年人口增长(GPOP)		-0.097 (0.212)			-0.161 (0.209)	
劳动力质量($QL1^*$)	0.108 (0.021)	0.104 (0.023)	0.076 (0.027)			
劳动力质量($QL2^*$)				0.094 (0.016)	0.090 (0.016)	0.072 (0.021)
收集到的评鉴分数(TEST)			-1.392 (1.455)			-0.628 (1.436)
受观察的劳动力质量(TEST $QL1^*$)			0.054 (0.032)			
受观察的劳动力质量(TEST $QL2^*$)						0.034 (0.028)
常数	-1.606 (0.749)	-1.184 (1.241)	-0.475 (1.069)	-1.483 (0.584)	-0.869 (0.984)	-0.657 (0.798)
国家数	78	78	78	80	80	80
R^2	0.41	0.42	0.49	0.41	0.41	0.48

注:系数下面括号内的数字是 Huber-White 标准误差。

表 5 的增长模型是表 2 增长基本模型的复制。表 5 的增长估计值同基线估计值很一致。让我们来考察一下因变量 $QL1^*$ 。纵栏(1)表明相关数据叙列具显著的有条件收敛性:亦即,初始收入的增加可以翻译为增长率的下降。学校教育数量变量具正符号,但不具统计显著性;再者,其业经估计的对增长的效应接近我们在表 2 所观察到的效应数值。最重要的是,劳动力质量(其值借 $QL1^*$ 测定)呈统计显著性,其数量等级与位于基线的国家样本组(此组仅含若干个国家样本)的估计值相同。质量变化 1 个标准差(1σ)可以翻译为年实际增长率的差异大约 1 个多百分比点。看起来,质量改进效应的重要性也大大超过数量增加效应:学校数量变化 1 个标准差可以翻译为平均增长 0.32 百分比点(即使这一点也可能因 Bils 和 Klenoy 的理论命题而受到夸大)。

表 5 的其他纵栏表明:受估计的质量效应具有稳定性。如在纵栏(2)加上人口增长率变量对受估计的劳动力质量效应,也并没有产生什么影响。纵栏(3)提供有关增长方程式(即方程式(1))中的劳动力质量变量的理论值。在求理论值时,我们先从具有劳动力质量的实测值的国家中将那些只有理论值的国家区分出来。对边缘测验分数效应($TESTQL1^*$)的点估计表明,那些具有直接观察样本的国家的测验影响稍微强些。虽说参数系数不具显著性(按 5% 水平)。正如前面所发现的那样,由备择选用的劳动力质量测度($QL2^*$)所得到的结果实际上是相同的。

按照 Levine 和 Renelt(1992 年)的研究结果,虽然加入许多其他国家经济的共同变量(在 Levine 和 Renelt 的论文中并未示明),但就一致性强弱和统计显著性而言,劳动力质量差异的重要性仍然不变。Barro(1991 年)和其他研究者们已经强调那些可能影响经济增长的不同政治因素。当我们将政治暗杀或者革命和政变的变量引入我们的模型,这些变量一致地不呈统计显著性。因此,我们并不削减劳动力质量差异的重要性。重要的是,下面关于受估计的质量效应的讨论产生这样一种可能性:劳动力质量变量可以部份地充当若干受省略的变量组的代理变量。但是,这种敏感度分析却将许多常用的、曾为人们使用过的备择变量摒弃了。

采用这种增广样本的估计值对扩大国家组这一类设想的适当性做出确认。这种扩大样本提高进一步检验假设的精度。它也指明,人们可以更广泛地将世界各国经济一体化。

5. 因果性, B 篇:劳动力质量 - 生产率关系 从省略变量的简单问题也会导致因果性的混淆。就一国的其他方面能对该国的学生成绩分数和该国经济成就的影响程度而言,劳动力质量变量可以简单地充当真实影响的代理变量。譬如说,如果增长同更加开放的劳力市场相关,那么,具有更开放的劳力市场的国家,其工人受教育的配置就更佳,其学生学习成绩也更佳。或者说,个人健康投资既可导致生产率和增长的升高,亦可导致学校教育业绩的改善。这类实例和其它实例会产生那种我们在此处所估计的劳动力质量 - 经济增长相关关系(即使在工人本身认知技能并不驱动经济增长的条件下)。举个例子来说,设想有这样的情况:一是共同因素(θ)影响一国的经济增长和学校教育效率,但受测定的劳动力质量变量(QL)并不进入该国的经济增长:

$$(5) g_i = f(\theta_i, X_i\beta) + \epsilon_i$$

$$(6) QL_i = g(\theta_i, Z_i\alpha) + u_i$$

这是那类省略变量问题的简化结构模型。根据假设,这类省略变量问题已经以某种形式或其他形式被引进到先前许多增长模型的估计和阐释程序了。没有变量 θ , 人们就很难处理这类情况。这是因为,一般说来,这种情况要求人们去寻觅那些与真实的学生和工人质量相关、但与增长毫无相关的工具变量。

在我们的分析里,我们采用一种不同于前人的策略。我们集中考察在美国工作的移民,并且将移民收益变异关联到我们的劳动力质量变量。如果我们的劳动力质量变量不表明生产率

差异,而且此变量本身仅仅是每一国家经济的其它差异的代理变量;那么,我们就应该认为:劳动力质量变量与美国国内经济的移民收入差异并无相关性。再者,我们也将移民受教育地点区分出来(移民在母国、在美国或者在母国和美国受教育)。因此,我们就有可能更密切地将劳动力质量变量和教育关联起来,而没将劳动力质量变量和移民特征关联起来。这些移民特征可能是文化的、家庭行为的特征或诸如此类的特征。

表6 学校教育质量对美国移民收入的影响(1989年*)

	只从具有劳动力质量测度观察值的国家移入的移民(n=37)			从具有劳动力质量测度观察值的国家或从具有劳动力质量测度预测度的国家移入的移民(n=87)					
	受教育所在地			总计		受教育所在地			
	只在母国受教育	在美国和母国受教育	只在美国受教育			只在母国受教育	在美国和母国受教育	只在美国受教育	
学校教育年数	0.089 (0.002)	0.090 (0.002)	0.132 (0.005)	0.103 (0.001)	0.099 (0.001)	0.089 (0.001)	0.117 (0.003)	0.130 (0.004)	
潜在经历	0.061 (0.002)	0.061 (0.002)	0.067 (0.008)	0.058 (0.002)	0.058 (0.002)	0.067 (0.002)	0.085 (0.003)	0.066 (0.006)	
(潜在经历) ²	-0.0009 (0.0000)	-0.0009 (0.0000)	-0.0016 (0.0001)	-0.0008 (0.0000)	-0.0008 (0.0000)	-0.0009 (0.0000)	-0.0015 (0.0001)	-0.0011 (0.0002)	
QL2*		0.0019 (0.0004)	0.0004 (0.0016)		0.0060 (0.0003)	0.0064 (0.0004)	-0.0025 (0.0006)	0.0018 (0.0011)	
常数	8.154 (0.035)	8.046 (0.041)	7.597 (0.127)	7.883 (0.022)	7.678 (0.024)	7.589 (0.031)	7.498 (0.054)	7.542 (0.093)	
观察样本	20,644	20,644	3,412	3,9840	3,9840	2,6643	8,639	4,558	
R ²	0.15	0.15	0.18	0.23	0.23	0.23	0.26	0.19	

注:a. 用以估计 Mincer 收入模型的因变量是劳动收入的对数。样本包括那些具有 1989 年劳动力市场收入的移民。这些移民出生的国家或者具有劳动力质量测度(QL2*)的观察值;或者,如本表右侧各栏所示,或者有观察值或者有估计值。

我们从《1990 年 5% 人口普查公众使用微型(光盘)文件》检取到所有海外出生男性工人的资料:在 1989 年时,年龄 50~60 岁;收入 1 000 美元以上。我们将这些资料构造出一个样本。《公众使用微型(光盘)文件》数据集提供劳动收入,上学年龄,出生国家,以及移居美国的年龄。我们构造出两个部分重叠的样本:(1) 那些在其出生国可以收集到其直接劳动力质量测定值(37 个国家)的移民工人;(2) 那些收集于增广样本(增广样本的被增加的国家的劳动力质量的测定值是估计值)[87 个国家]的移民工人。用常用最小二乘法估计出来的收入方程式以标准 Mincer 型方程式开始[参阅 Mincer:《学校教育,经历和收入》(1974 年)],在标准 Mincer 型收入方程式里,共回归倚变量是年收入的 \ln ,其回归自变量是:学校教育年数、潜在经历(年龄-学校教育-6)变量和潜在经历变量的平方;再加上我们以出生国为基础的劳动力质量($QL2^*$)变量。

表 6 所示的男性移民的简单 Mincer 模型(在表中两个样本的每一样本的纵栏(1)提供相当标准的结果:收入梯度(每学年 9~6%)和每次跳槽的报酬递减。饶有兴趣的是受估计的劳动力质量效应。在受观察的样本中,测验分数 1 点可以翻译为平均个人收入升高 0.19%。在表中,上述纵栏之后的三个纵栏是按移民者受教育的国别分类的分类样本(以移民者移入美国的年龄为依据)。那些完全在母国受教育的移民,其测验分数每升高 1 点,平均收入提高 0.21%。那些在美国受到部份教育或全部教育的移民,其个人和母国劳动力质量($QL2^*$)间的相关呈统计不显著性。至于在该表(表 6)右侧的增广样本,其计算结果的模式并无不同。并且,估计得到的劳动力质量对个人收入的效应甚至还要强些。

如果先前估计的劳动力质量对增长的效应只反映若干个业经省略的、与母国及母国市场有效性相关的因子,那么,我们就不应该期望能弄清移民劳动力质量对美国经济增长的效应。再者,如果增长效应只单纯地反映文化因子或者家庭因子,并不反映学校教育成绩和技能效应,那么,我们就可以期望劳动力质量对美国经济增长的效应仍然是不变的;这时,我们也不必考虑移民究竟是否在母国或者是在美国受到学校教育的。我们对表 6 所得的结果的解释是,此结果提供这样一种一致性的证据:我们模型的劳动力质量的变量是与个人的生产能力相关的;并且表明它对经济增长有因果性影响。

不管如何,只要这种劳动力质量论证支持国别劳动力质量差异与个人能力之间的因果联系性,问题仍然是:受估计效应的等级是充分到足以解释劳动质量与增长率间具有强相关性。从人们先前从那些对学校资源和经济增长的相关性可能并非时序的调研得到的结论是:劳动力质量影响的途径与数据并不一致。但是,这种调研和其后续的、与东亚国家相关的劳动力质量对增长效应差异(参阅下节)的调研只是质的问题,它们对受估计的质量—增长效应等级并没作出什么说明。关于不同学校教育引致的技能差异和生产率等级异差的问题,个人收入估计工作对此给出直接论证。因此,这种工作也使人们有可能去对增长效应等级问题作些说明。

不幸的是,眼下并没有简单易行的、能将个人生产率效应(这种效应与收入水平有关)翻译成增长总效应的方法。许多从事这种翻译的方法取决于人们所选择的特殊增长模型。譬如说,一种方法是企图这样来将那种能影响某国经济体的稳态总产出水平的教育成就效应模型化。这类经济体的总产出水平是从初始收入通过条件收敛而生成的。在这种模型中,质量系数的相对等级和基于初始收入的系数给出质量对稳态收入水平的影响估计值。稳态收入水平是这样一个参数:只要增长是直接源于人力资本投入水平的变化,该参数就应该可以和 Mincer 收入参数相比较。然而,此增长估计值提供出比人们在移民收入估计方面所得到的数值还要大得多的增长估计值。这种现象使人们作出这样设想:比简单而直接的影响还大的生产率效应正“渗透”增长相关性。

另一个源于内生增长模型视角的观点将增长率和人力资本(或者按量而言,或者按质而

言)的存量关联起来。根据这种观点,强外部性或者用劳动力质量总存量表达的内生性增长可能有能力来解释那种用个人生产率和劳动力质量差异的增长效应所表述的明显不对称性。在查定这种明显不对称时有两点很重要。因为在增长方程式中,劳动力质量系数非常大(不论就绝对值或者就其对学校教育数量的相对值而言),人们就必需强调在那种与数量外部性不同的质量外部性的强度。其次,在一个内生性增长模型实例中,劳动力质量效应等级仍然是难以置信地大。在基线增长模型和增广增长模型的劳动力质量估计值表明:测验分数的一个标准差可以翻译为年平均人均 GDP 增长率升高 1% 多一点。但是,根据 Klenow 和 Andres Rodriques - Claire 的估计值,大调查组的跨国间平均增长(这种增长可归功于技术改进)稍微超过 1% (年度 GDP 增长率);粗略地说来,此增长值与测验分数的一个标准差相应。这样,即使来自其对技术变化的总影响,受估计的劳动力质量对经济增长的效应也显得太大了。

其它模型既允许劳动力质量影响收入水平,也允许劳动力质量影响总增长率。这种“穿插式”影响方式将影响效应的等级差异调和起来。虽说如此,但我们却得到这样的结论:眼下,由于测定的劳动力质量的直接因果性影响所造成的受估计增长效应究竟有多大还没有定论。其中一部份效应可能来自那些与测定劳动力质量有强相关关系的、被省略的变量。然而,这些被省略掉的变量的精确性质还不清楚。这是因为敏感度分析和其它因果结构将许多颇有说服力的因子排除了。这样,我们作出这样结论:有关这类问题的论证表明,劳动力质量变量确实显示出生产率差异。此类论证也作出确认:和文化的、人种的或者父母的因素相反,在影响劳动力质量差异方面,存在着跨国间学校教育的角色作用。不过,由于微观效应似乎被直接翻译为较小的增长影响,目前关于完整的因果结构问题仍然还没有解决。

表 7 东亚国家的重要性

	具有劳动力质量测度观察值的国家				具有劳动力质量测度预估值国家			
	整个样本	不包括四小虎 ^④	不包括高绩效国家 ^⑤	不包括新兴工业化国家 ^⑥	整个样本	不包括四小虎 ^④	不包括高绩效国家 ^⑤	不包括新兴工业化国家 ^⑥
初始人均收入(Y60)	-0.472 (0.090)	-0.357 (0.092)	-0.328 (0.104)	-0.270 (0.086)	-0.382 (0.078)	-0.308 (0.081)	-0.291 (0.084)	-0.256 (0.086)
学校教育数量(S)	0.103 (0.118)	0.117 (0.116)	0.106 (0.118)	0.085 (0.113)	0.127 (0.086)	0.148 (0.087)	0.140 (0.089)	0.143 (0.091)
劳动力质量(QLI*)	0.134 (0.020)	0.101 (0.022)	0.095 (0.024)	0.091 (0.023)	0.108 (0.020)	0.079 (0.020)	0.075 (0.020)	0.069 (0.020)
常数	-1.901 (0.936)	-1.111 (0.934)	-0.929 (1.010)	-0.966 (0.986)	-1.606 (0.730)	-0.848 (0.719)	-0.709 (0.739)	-0.648 (0.755)
国家数	31	27	26	25	78	74	73	70
R ² (调整过)	0.70	0.49	0.39	0.40	0.39	0.26	0.22	0.21

注:系数下面括号内的数字是 Huber - White 标准误差

④ 四小虎:香港地区、韩国、新加坡和台湾地区。

⑤ 高绩效国家:香港地区、韩国、新加坡和台湾地区加上日本

⑥ 新兴工业化国家:香港地区、韩国、新加坡、台湾地区和日本加上印尼,马来西亚和泰国(印尼和马来西亚都没有参加任何一次国际测验)。

6. 因果性, C 篇: 模型对东亚国家的敏感度 在我们的样本时期里, 东亚国家的举世闻名的经济经历引起人们注意到该地区国家基本上已经取得高经济增长成果。正如所看到的那样, 在各种国际测验方面, 东亚国家有一种赢得高分的倾向。因此, 即使在劳动力质量和经济增长并没有因果关系的情况下, 国际测验分数也可能只是对东亚国家的教育一种评分。人们可以将这种特殊情况想象为上一节所述的、对被省略的变量的一种特例。

表 7 证明这些国家劳动力质量对增长因变量的影响相关性。该表将整个样本的计算结果和扣除各个东亚国家的不同小组的余留国家的计算结果加以比较。我们探讨三个东亚国家(地区)小组: 四小虎(香港地区、韩国、新加坡和台湾地区); 高增长业绩国家(地区)(四小虎 + 日本)和新兴工业化国家(地区)(高增长业绩国家 + 印尼、马来西亚和泰国)。我们就那个只包括具有测验分数实测值国家的样本和那个增广样本(即包括那些具有预期测验分数的国家), 对这些模型求估计值。就各种估计值的情况而言, 初始收入水平的负效应应呈显著性。这表明条件收敛。虽说在包括东亚国家以后, 这些负效应的强度还要大些。当大样本减掉东亚国家之后, 学校教育数量的估计效应强度稍微大些; 当将样本只限定于那些具备测验分数实测值时, 受估计的学校教育数量效应强度并不显示相应变化; 不管是那种情况, 学校教育数量效应变量都是不呈统计显著性的。

就具有测验观测值的样本而言, 劳动力质量效应降级(1/3 ~ 1/4)的跌落程度随样本的精确度而异; 在删除东亚国家小组之后, R^2 大幅度跌落。这些结果和那种对东亚国家经济增长作出贡献的人力资本质量相一致。不管如何, 劳动力质量对各个国家小组仍然具有统计显著性的强度效应。就增广样本而言, 这种结果仍然有效。

简单说来, 劳动力质量的重要性并不仅仅是由于样本扩及东亚国家致使统计数据膨胀而形成的。就全世界的各个地区而言, 劳动力质量仍然具有重要性。在东亚以外的地区里, 劳动力质量变化 1 个标准差仍然意味着增长率升高 0.7 ~ 1.0 百分比点(视用于估计的特定样本而异)。

7. 劳动力质量测度的重要性 表 8 给出具质量测度的估计模型对不具质量测度的估计模型(或具学校教育投入测度的估计模型)的直接比较。无质量测度(纵栏(1) ~ 纵栏(3)), 学校数量测度呈一致显著性。再者, 从纵栏(2), 小学学生 - 教员比率测度虽呈不显著性(在 10% 水平), 但等级小。将小学学生 - 教员比率平均值削减 50%(从当年世界平均值削减 18)(这种大大增加教员的措施是一项非常昂贵的教育政策)会使经济增长率升高 0.7%。就中等学校而言, 学生 - 教员比率的效应反常, 但不具统计显著性。不过, 我们可以认为, 每名学生的费用总额并不给经济增长多带来什么额外的信息。

然而, 劳动力质量的直接测度却有非常不同的影响(纵栏(4)和纵栏(5))。劳动力质量变量的方差份额增加了一倍而达到 40%。此外, 在表明人力资本对国民经济增长的重要性上, 人力资本的质量测度取代了数量测度。

表 8 的纵栏(6)和纵栏(7)分别是质量的两个直接测度(纵栏(4)和纵栏(5))同投入测度(纵栏(2)和纵栏(3))的组合(纵栏(6) = 纵栏(2) + 纵栏(4); 纵栏(7) = 纵栏(3) + 纵栏(5))。在对作为劳动力质量测度的纵栏(4)或纵栏(5)作条件处理之后, 纵栏(2)和纵栏(3)的投入均不对增长有显著的影响。纵栏(6)和纵栏(7)这两个测度只进一步肯定这个判断: 与学校教育投入测度相比较, 测验分数测度含有更多的信息量。

表 8 学校质量和增长的备择测度之比较

	基线	学校投入		质量测度		业已组合的投入和质量	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
初始人均收入(Y60)(100美元)	-0.407 (0.139)	-0.455 (0.114)	-0.408 (0.137)	-0.382 (0.081)	-0.370 (0.084)	-0.393 (0.095)	-0.368 (0.095)
学校教育数量(S)	0.529 (0.122)	0.481 (0.124)	0.486 (0.133)	0.127 (0.089)	0.120 (0.096)	0.070 (0.104)	0.065 (0.117)
小学学生-教员比率(PT-pri)		-0.040 (0.024)				0.001 (0.026)	0.006 (0.024)
中学学生-教员比率(PT-Sec)		0.027 (0.042)				-0.038 (0.044)	-0.038 (0.045)
教育总支出/GDP(EXPEND)			14.42 (14.73)			7.388 (16.060)	3.968 (15.100)
劳动力质量(QL1*)				0.108 (0.021)		0.112 (0.020)	
劳动力质量(QL2*)					0.094 (0.016)		0.100 (0.015)
常数	0.933 (0.238)	2.140 (2.060)	0.558 (0.418)	-1.606 (0.749)	-1.483 (0.584)	-1.113 (1.091)	-1.042 (0.992)
国家数	100	96	96	78	80	76	78
R ²	0.23	0.26	0.22	0.41	0.41	0.42	0.42

注:系数下面括号内的数字是 Huber-White 标准误差。

表 9 备择增长模型的均方根误差

模型(由表 8 整理而得)	各国	平均增长率国家组		
		< 1%	1 ~ 3.5%	> 3.5%
A. QL1* 质量测度				
基线: C, Y60, S	1.50	2.16	0.87	2.16
投入: C, Y60, S PT - pri PT - sec	1.50	2.16	0.87	2.17
投入: C, Y60, S EXPEND	1.50	2.14	0.86	2.18
质量: C, Y60, S QL1*	1.28	1.95	0.79	1.71
国家数	76	13	47	16
B. QL2* 质量测度				
基线: C, Y60, S	1.52	2.08	0.89	2.21
投入: C, Y60, S PT - pri PT - sec	1.52	2.07	0.89	2.22
投入: C, Y60, S EXPEND	1.52	2.07	0.88	2.23
质量: C, Y60, S QL2*	1.31	1.90	0.87	1.67
国家数	78	15	47	16

在表 9 中,合适的质量测度的重要性令人一看就明白。表 9 对那些用于解释一个具有一致性的国家组的增长率的均方根误差作了比较。在表(8)中,与其它模型相比较,质量测度模型的均方根误差下降 0.2 百分比点(按年度增长)。再者,在解释增长率的不规则性方面,质量测度特别有效。表 8 最后三个纵栏将各国分为慢增长国家(年平均增长率低于 1%)、快速增长国家(年平均增长率高于 3.5%)和其他国家。增长最快的几个国家和增长最慢的几个国家的均方根误差都出现下跌,特别是那些用 QL1* 能更好说明其增长迅速的国家,其均方根误差的跌幅更突出。在用 QL1* 去解释增长差异时,快速增长国家的均方根误差的跌幅大约是 2.2% 至 1.7%。就绝对数量而言,增长分布的底线国家的均方根误差跌幅较小。虽说如此,但跌幅还是可观(为

0.2%)的。

图 2 国家分布条形图(按实际人均 GDP 的年平均增长率分)

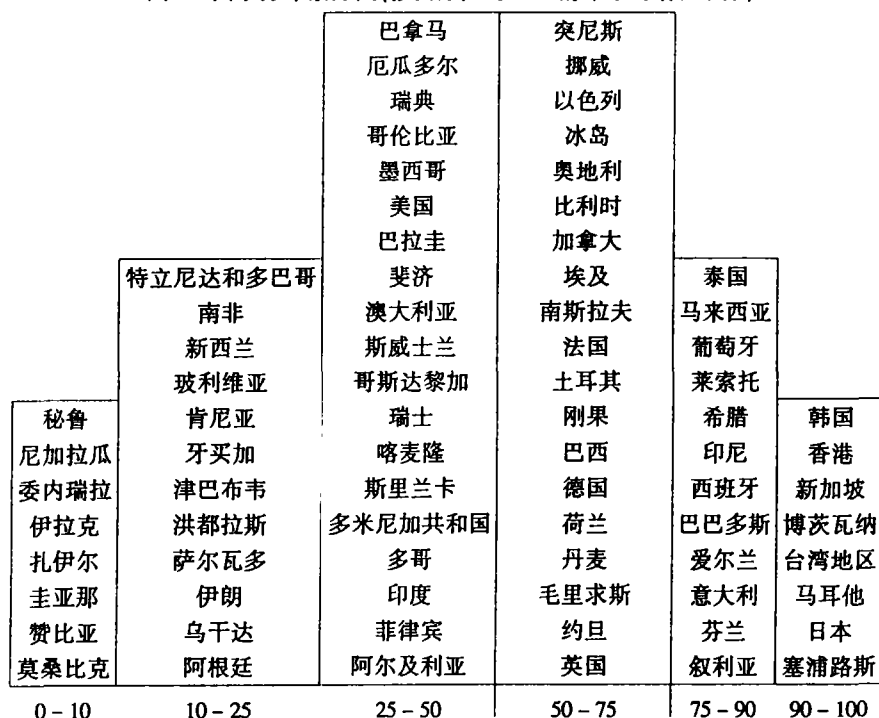
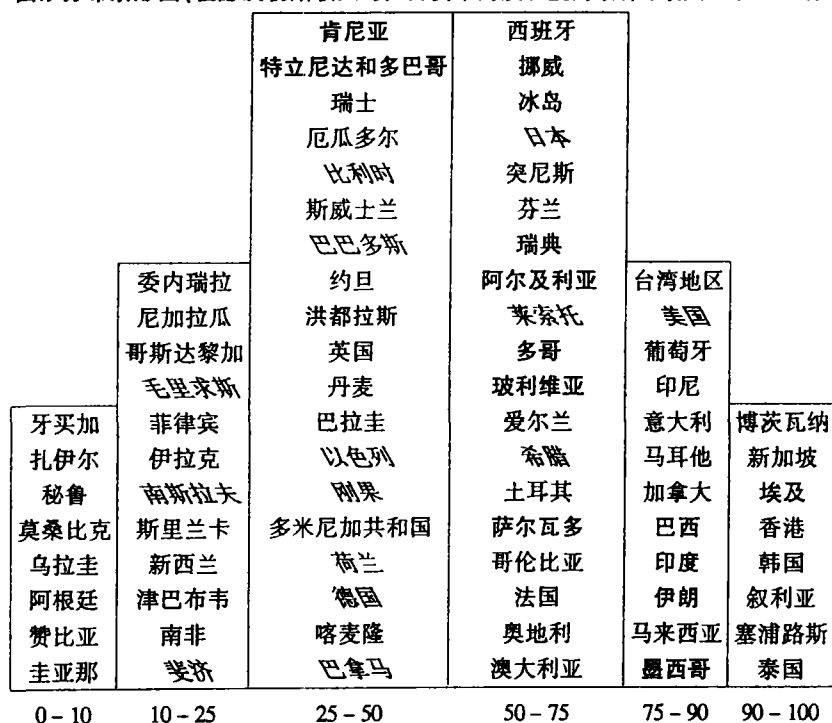


图 3 国家分布条形图(在虑及初始收入与人力资本的质和量的条件下,按人均 GDP 增长率分)



注:用黑体字书写的国家(肯尼亚、特立尼达和多巴哥、洪都拉斯、阿尔及利亚、多哥、玻利维亚、萨尔瓦多、美国、巴西、印度、伊朗、墨西哥和埃及)是其在条件分布下比在非条件分布下高 20 或 20 个百分点的国家。用斜体字书写的国家(毛里求斯、南斯拉夫、斐济、比利时、巴巴多斯、比利时、刚果、荷兰、德国、巴拿马、日本、莱索托和希腊)是其在条件分布下比在非条件分布下低 20 或 20 个百分点的国家。

先前的分析集中在那些可以用初始条件和人力资本质量和数量来解释的经济增长率差异。同时,这些高度简化的模型将广泛的国家经济特征舍弃了。受观察增长率分布反映“增长条件”和不受测定的因子的组合,借助表9所列的分析,人们已经弄清楚:跨国间的受测因子和不受测因子的重要性是不同的。在舍弃受测效应,我们可以在给定的人力资本存量条件下识别哪些国家的增长是快得惊人,或者哪些国家是慢得惊人。图2列示出按1960年~1990年年平均增长率归类的国家百分位数的分布情况。图3列示出按初始收入和人力资本数量和质量条件处理了的国家百分位数的分布情况。也就是说,图3列示出那些其增长业绩比人们期望的业绩好得惊人和坏得惊人的国家。图3用黑体字表明的国家有13个。就其分数分布情况而言,这13个国家的百分比分的排序都是按这样的原则:其有条件的增长至少比无条件的增长高出20点。譬如埃及,其平均增长率为2.9。从按无条件的增长排序的第27位国移到其人力资本存量条件的增长排序的第3位国。在同一时期内,日本的平均增长率为5.3%,从按无条件的增长排序的第7位国下降到第24位国(在扣除其有利条件的增长之后)。

在消除受测的初始收入到人力资本的影响之后,增长率的分布给人们提供一条去寻求、从而去识别更多影响增长因子的线索。譬如说美国,人们在给定人力资本特征条件下认定其增长比预期速度快得多(从国家排序中的低于中位数升到列入最高10百分数的国家组中)。或许,这种现象又部份地反映出人力资本测度的不完备性。这种分析和美国的高等教育报导完全沾不上边。因为在高等教育方面,美国的高校是世界最佳的高校的说法颇有争议。再者,在那些强调观念 and 创新的内生性增长模型里,人们将较高的教育投入物投放在科学家和工程师的生产是特别重要的。另一种说法是,或许美国增长比预期还快只简简单单地反映了美国的市场更开放、更有竞争性而已。

饶有趣味的是将这种看法装到经常引人注目的东亚国家经济增长的框架上去。既然考虑到劳动力质量测度,人们也就将新加坡、香港、韩国、泰国和台湾视为增长非常迅速的国家 and 地区了。换句话说,东亚经济奇迹包含某种引人注目的成份。这样一种成份已经超出对人力资本开发的任何强调和人力资本开放的实际成就。与此同时,日本增长模型的非凡性成就似乎就差多了。这种说法不是要极力贬低人力资本对发展的重要性。重要的是人们应该注意到某种引人注目的增长成份是在别的地方!

8. 结论 我们已经将各国经济增长率看遍。最稳固和最易于为人接受的一个结论还是摆脱不了个别国家人力资本这个讨论中心。然而,这种结论却是来自那些人力资本变量很不一致的设定的模型。实际上,所有这类模型都无视质量问题。它们只有隐含地假设:相对于人力资本的纯数量差异和重要性,人力资本的质量差异是很小的。

本文的分析凸显地将劳动力质量视为一种用数科技能比较测验来计量的变量。将该变量的不同分析加以界定,我们就可以得到单一性结论:劳动力质量具有一种对经济增长的一致而稳定的强相关关系。就数量而言,对模型基础性的变量规定的一系列间接调研是与因果阐释相一致的。看来,增长相关关系并不是这样一种特殊增长的后果:这种增长将资源投在学校来提高劳动力质量;看来,增长相关关系也不是受高测验得分所驱动的,这种高得分只简单地与东亚国家相关。因为就其他理由说来,这些东亚国家也可以得到高增长。最后,在考察我们所作的劳动力质量变量和美国移民收入的相关性情况时,我们找到了昭然若揭的证明:国际测验分数和生产率差异有相关关系。看来,这些生产率差异是和学校教育差异相关的,而不是和文化因素、家庭支持和态度,等等相关的。这种和生产率的直接关联表明:劳动力质量对国际经济业绩有因果性影响。

受估计的质量对增长的影响表明:数科技能的1个标准差可以翻译成为年平均实际增长提高1个多百分比点。因此,看来,这种影响之大,令人难信。目前并没有一种意见一致的简

单方法能将那种由个人收入引起的生产率差异翻译成经济增长效应。当然,这种简单方法也使得这种翻译难于精确。譬如说,如果劳动力质量由于收入水平稳定提高而越来越发挥作用,如果经济增长反映一种条件收敛过程,那么,从增长方程式求得的经济增长的结果就应该大大地超过个人收入的增长结果。此简单方法以外还有一种方法。这种方法的模型与本文所提出的具有基础性假设的内生性增长模型更一致。此模型方法将人力资本水平的提高所生成的外部条件的角色作用彰显出来。然而,从增长模型求得的结果却暗示,质量性外部条件应该比数量性外部条件强得多。据估计,劳动力质量的 1 个标准差对增长的效应比 9 年平均学校教育所发挥的对经济增长的效应还要大。再者,按绝对数量来说,这种效应大约等于这时期平均技术进步率的估计值。这些命题表明:在增长方程式中,可能存有被省略的因子。但是,在本文中,我们并不明指这些被省略的因子是些什么因子。由于变量规定的分析、由于对另一种因果效应模型的考虑,许多仿佛存在的因子被舍弃了。

我们得出这样的结论:对经济增长而言,劳动力质量差异是重要的,这类质量差异是与学校教育相关的(但不一定是与一个国家投资教育的资源相关的);劳动力质量有一种对增长的因果性影响。同时,跨国间增长相关性的简单估计值似乎夸大劳动力质量的因果性影响。目前,这种夸大的精确原因或夸大程度尚未弄清楚。

显然,在这里还出现政策的两难问题。之所以如此,其原因是标准学校资源政策与那些业已识别的劳动力质量差异并不相关。我们相信,对导致劳动力质量差异的基础性因子的进一步考察是很重要的。

附录 A 变量和资料来源表

(BL) Barro 和 Lee:《教育成就和国际比较》,载《货币经济学杂志》,1993 年 12 月。(附 1994 年新资料)。

(ER) Easterly 和 Rebelo:《财政政策 and 经济增长:一种经验研究》,载《货币经济学杂志》,1993 年 12 月。

(KL) King 和 Levine:《金融,企业主和增长:理论和论证》,载《货币经济学杂志》,1993 年 12 月。

(SH) Summers 和 Heston:《Penn 世界表(第 5 号):1950 年~1988 年增订国际比较表》,载《经济学季刊》,1991 年 5 月。

(联合国教科文组织) 联合国教科文组织出版(巴黎)《统计年鉴》,各年。

分析模型的变量

变量 Asia, Latin America 和 Africa: 亚洲国家、拉美国家和撒哈拉沙漠以南非洲国家的地区虚拟变量

(BL 文献:变量 ASIAE, LAAM 和 SAFRICA)

变量 GCNL:用于国防和教育的政府实际支出净额对实际 GDP 的比率

(ER 文献:变量 HSGRXDXE)

变量 ENROLL - pri:初级教育毛入学率总数的平均值

(BL 文献:变量 Pxx)

变量 EXPEND:政府对教育的名义支出占名义 GDP 的比率的平均值。

(BL 文献:变量 GEETOT)

变量 GPOP:人口增长率的平均值

(BL 文献:变量 GPOP)

变量 GR:1960 年~1990 年时期的人均 GDP 平均增长率。如果无法弄到 1960 年至 1990 年的资料,则采用可以得到的子时期的资料。变量人均实际 GDP 则 SH 文献的变量 RGDPC (按 1985

年国际价格计算)。

变量 PPE: 每名学生当年公共支出, 即 $GDP^* (POP)^* (当年支出/GDP)/(中学学生人数 + 小学学生人数)$ [单位 100 美元] (SH 文献: 变量 GDP; 联合国教科文组织文献: 变量当年支出, 变量小学学生人数和变量中学学生人数)

变量 PRINV: 私人投资占 GDP 的比率的平均值。

(BL 文献: 变量 INVWB - INVPUB)

变量 PT - pri.: 平均小学学生/教师比率

(BL 文献: 变量 TEAPR1)

变量 PT - sec.: 平均中学学生/教师比率

(BL 文献: 变量 TEASEC)

变量 QL1、QL2: 本文正文的学校教育质量的测度 (参阅下述)。

变量 QL1*, QL2* 上述质量测度的扩大测度 (包括国际学生成就测验比较调研的非参与国和参与国。非参与国采用预估值)。参阅本文正文及下述。

变量 RECUR: 当年名义政府教育支出占名义 GDP 的比率

(BL 文献: 变量 GEEREC)

变量 S: 1960 年, 1965 年, 1970 年, 1975 年, 1980 年和 1985 年学校教育年数的算术平均值。

(BL 文献: 变量 HUMAN)

变量 TEST: 学生成绩测验国际比较调研的参与国用的虚拟变量。(参阅本文正文)。

变量 TRD: 总贸易量/GDP (文献 KL)。

变量 Y60: 初始收入, 1960 年 (单位: 1000 美元)

(SH 文献: 变量 EGD, PCH)

附录 B 用于估计生产函数的资料

利用一个 (参阅表 3) 生产函数估计值将不同测验年份联系起来的样本。此样本的因变量是每一个国家和每一次国际测验的平均标准化分数。下面表 B1 记述生产函数模型的外生变量的日期。日期的计算基础是学生成绩测度的特定测验年份。

表 B1 记载生产模型外生变量的国际教育组织测验日期

变量名称	IEA ^① 数学 测验日期 1964 年 ~ 1966 年	IEA ^① 科学 测验日期 1966 年 ~ 1973 年	IEA ^① 数学 测验日期 1980 年 ~ 1982 年	IEA ^① 科学 测验日期 1983 年 ~ 1986 年	IAEP ^② 1968 年	IAEP ^② 1991 年
S_{t-1}	1960	1960 - 1965	1965 - 1975	1970 - 1980	1975 - 1985	1975 - 1985
$EXPEND_{t-1}$	1960 - 1964	1960 - 1969	1965 - 1974	1970 - 1979	1975 - 1984	1975 - 1984
$PT - pri_{t-1}$	1950 - 1960	1955 - 1965	1960 - 1970	1965 - 1975	1970 - 1980	1970 - 1980
PRE_{t-1}	1960	1960 - 1965	1965 - 1975	1970 - 1980	1975 - 1985	1975 - 1985
$GPOP_{t-1}$	1960 - 1964	1960 - 1969	1965 - 1974	1970 - 1979	1975 - 1984	1975 - 1984

注: ① 为“国际教育成就计量学会”

② 为“国际教育进步评鉴组织。”

附录 C: 劳动力质量的数据

下列数据 (详见表 (1)) 是依本文正文所述各种方法生成的。

虚拟变量 TEST = 1, 当学生成绩分数属实际测验值时, QL1* 和 QL2* 都是受观测的测验数据, 若虚拟变量 TEST = 0 则 QL1* 和 QL2* 都是按模型计算得到的数据。若虚拟变量阙如, 则要么用于测算学生成绩分数所需的资料或者用于估计增长模型所需的数据的阙, 要么测算得到的学生成绩分数低于 20 (参阅正文)。

表 C1 劳动力质量数据

国家和地区	虚拟变量 TEST	劳动质量 QL1 *	劳动力质量 QL2 *	国家和地区	虚拟变量 TEST	劳动质量 QL1 *	劳动力质量 QL2 *
1. 阿尔及利亚	0	28.28	28.06	35. 多米尼加	0		
2. 安哥拉	0			36. 多米尼加共和国	0	37.41	39.34
3. 阿根廷	0	42.99	48.50	37. 厄瓜多尔	0	35.78	38.99
4. 澳大利亚	1	48.13	59.04	38. 埃及	0	26.01	26.43
5. 奥地利	0	53.20	56.61	39. 萨尔瓦多	0	21.81	26.21
6. 巴哈马	0			40. 埃塞俄比亚	0		
7. 巴林	0	26.03	23.19	41. 斐济	0	50.02	58.10
8. 孟加拉国	0			42. 芬兰	1	48.76	59.55
9. 巴巴多斯	0	50.41	59.80	43. 法国	1	54.15	56.00
10. 比利时	1	53.25	57.08	44. 加蓬	0		
11. 伯利兹	0			45. 冈比亚	0		
12. 贝宁	0			46. 西德	1	59.03	48.68
13. 布丹	0			47. 加纳	0		25.58
14. 玻利维亚	0	22.10	27.47	48. 希腊	0	49.11	50.88
15. 博茨瓦纳	0	25.05	31.71	49. 格林纳达	0		
16. 巴西	1	33.91	36.60	50. 危地马拉	0		
17. 保加利亚	0			51. 几内亚比绍	0		
18. 布基纳法索	0			52. 几内亚	0		
19. 布迪隆	0			53. 圭亚那	0	45.71	51.49
20. 喀麦隆	0	39.00	42.36	54. 海地	0		
21. 加拿大	1	47.57	54.58	55. 洪都拉斯	0	26.43	28.59
22. 佛得角	0			56. 香港	1	56.93	71.85
23. 中非共和国	0		24.77	57. 匈牙利	1	53.85	61.23
24. 乍得	0			58. 冰岛	0	48.13	51.20
25. 智利	1	26.30	24.74	59. 印度	1	21.63	20.80
26. 中国	1	59.28	64.42	60. 印尼	0	37.98	42.99
27. 哥伦比亚	0	34.78	37.87	61. 伊朗	1	20.79	18.26
28. 科摩罗	0			62. 伊拉克	0	29.34	27.50
29. 刚果	0	46.04	50.90	63. 爱尔兰	1	47.59	50.20
30. 哥斯达黎加	0	42.15	46.15	64. 以色列	1	51.29	54.46
31. 塞浦路斯	0	42.24	46.24	65. 意大利	1	44.50	49.41
32. 捷克斯洛伐克	0			66. 象牙海岸	0		
33. 丹麦	0	53.48	61.76	67. 牙买加	0	44.19	48.62
34. 吉布提	0			68. 日本	1	60.65	65.50

续上表

国家和地区	虚拟变量 TEST	劳动质量 QL1 [*]	劳动力质量 QL2 [*]	国家和地区	虚拟变量 TEST	劳动质量 QL1 [*]	劳动力质量 QL2 [*]
69. 约旦	1	39.38	42.28	110. 卢旺达	0		
70. 肯尼亚	0	24.43	29.73	111. 沙特阿拉伯	0		
71. 大韩民国	1	56.21	58.55	112. 塞内加尔	0		
72. 科威特	0	28.36	22.50	113. 塞舌尔	0		
73. 老挝	0			114. 塞拉利昂	0		
74. 莱索托	0	46.14	51.95	115. 新加坡	1	56.51	72.13
75. 利比里亚	0			116. 所罗门群岛	0		
76. 卢森堡	1	39.45	44.49	117. 索马里	0		
77. 马达加斯加	0			118. 南非	0	45.25	51.30
78. 马拉维	0			119. 西班牙	1	49.40	51.92
79. 马来西亚	0	47.89	54.29	120. 斯里兰卡	0	41.54	42.57
80. 马里	0			121. 圣卢西亚	0		
81. 马尔他	0	53.16	57.14	122. 圣文森特岛	0		
82. 毛里塔尼亚	0			123. 苏丹	0		
83. 毛里求斯	0	49.53	54.95	124. 苏里南	0		
84. 墨西哥	0	35.06	37.24	125. 斯威士兰	1	35.46	40.26
85. 蒙古	0			126. 瑞典	1	47.41	57.43
86. 摩洛哥	0			127. 瑞士	1	57.17	61.37
87. 莫桑比克	1	24.26	27.94	128. 叙利亚	0	31.66	30.23
88. 缅甸	0			129. 台湾地区	1	56.28	56.31
89. 纳米比亚	0			130. 坦桑尼亚	0		
90. 尼泊尔	0			131. 泰国	1	39.83	46.26
91. 荷兰	1	56.84	54.52	132. 多哥	0	28.08	32.69
92. 新西兰	1	52.44	67.06	133. 汤加	0		
93. 尼加拉瓜	0	24.19	27.30	134. 特立尼达和多巴哥	0	40.57	46.43
94. 尼日利亚	1	34.25	38.90	135. 突尼斯	0	41.79	40.50
95. 尼日尔	0			136. 土耳其	0	41.52	39.72
96. 挪威	1	49.60	64.56	137. 乌干达	0		
97. 阿曼	0			138. 乌拉圭	0	46.33	52.27
98. 巴基斯坦	0			139. 阿拉伯联合酋长国	0		
99. 巴拿马	0	42.02	46.78	140. 英国	1	53.98	62.52
100. 巴布亚新几内亚	0		22.58	141. 美国	1	43.43	46.77
101. 巴拉圭	0	37.99	39.96	142. 苏联	1	53.89	54.65
102. 秘鲁	0	37.83	41.18	143. 瓦努阿图	0		
103. 菲律宾	1	34.35	33.54	144. 委内瑞拉	0	36.78	39.08
104. 波兰	1	50.28	64.37	145. 西萨摩亚	0		
105. 葡萄牙	1	44.09	44.22	146. 也门	0		
106. 波多黎各	0			147. 南斯拉夫	0	50.91	53.97
107. 卡塔尔	0			148. 扎伊尔	0	30.03	33.53
108. 留尼汪岛	0			149. 赞比亚	0	30.54	36.61
109. 罗马尼亚	0			150. 津巴布韦	0	35.97	39.64

译者单位:厦门大学南洋研究院

邮 编:361005